IFW16

```
 3 <110> APPLICANT: INOUYE, MASAYORI
 4        PHADTARE, SANGITA
 5        YAMANAKA, KUNITOSHI
 6        KATO, IKUNOSHIN
 8 <120> TITLE OF INVENTION: GENE ENCODING A 4,5 DIHYDROXY-2-CYCLOPENTEN-1-ONE
 9        (DHCP) EFFLUX PROTEIN PROMOTING RESISTANCE TO DHCP
11 <130> FILE REFERENCE: 1137-R-00
13 <140> CURRENT APPLICATION NUMBER: 09/805,681B
14 <141> CURRENT FILING DATE: 2001-03-14
16 <150> PRIOR APPLICATION NUMBER: 60/228,727
17 <151> PRIOR FILING DATE: 2000-08-29
19 <160> NUMBER OF SEQ ID NOS: 12
21 <170> SOFTWARE: PatentIn Ver. 2.1
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 3900
25 <212> TYPE: DNA
26 <213> ORGANISM: Escherichia coli
28 <400> SEQUENCE: 1
29 gccagccact cttccagctg acgcacggta tagctgaccg cagaaggaac gcgatgcagc 60
30 tcctgtgccg cagcgctaaa actaccatta cgcgctaccg catcaacaac ttcgagtgaa 120
31 tattctgacc acatagtctg cctgcaaaat ttttgaaacc agtcatcaaa tattaccgtt 180
32 tcacaacact aatttcactc cctacacttt gcggcggtgt ttaattgaga gatttagaga 240
33 atatacatgc aacctgggaa aagattttta gtctggctgg cgggtttgag cgtactcggt 300
34 tttctggcaa ccgatatgta tctgcctgct ttcgccgcca tacaggccga cctgcaaacg 360
35 cctgcgtctg ctgtcagtgc cagccttagt ctgttccttg ccggttttgc cgcagcccag 420
36 cttctgtggg ggccgctctc cgaccgttat ggtcgtaaac cggtattatt aatcggcctg 480
37 acaatttttg cgttaggtag tctgggggatg ctgtgggtag aaaacgccgc tacgctgctg 540
38 gtattgcgtt ttgtacaggc tgtgggtgtc tgcgccgcgg cggttatctg gcaagcatta 600
39 gtgacagatt attatccttc acagaaagtt aaccgtattt ttgcggccat catgccgctg 660
40 gtgggtctat ctccggcact ggctcctctg ttaggaagct ggctgctggt ccattttttcc 720
41 tggcaggcga ttttcgccac cctgtttgcc attaccgtgg tgctgattct gcctattttc 780
42 tggctcaaac ccacgacgaa ggcccgtaac aatagtcagg atggtctgac ctttaccgac 840
43 ctgctacgtt ctaaaaccta tcgcggcaac gtgctgatat acgcagcctg ttcagccagt 900
44 ttttttgcat ggctgaccgg ttcaccgttc atccttagtg aaatgggcta cagcccggca 960
45 gttattggtt taagttatgt cccgcaaact atcgcgtttc tgattggtgg ttatggctgt 1020
46 cgcgccgcgc tgcagaaatg gcaaggcaag cagttattac cgtggttgct ggtgctgttt 1080
47 gctgtcagcg tcattgcgac ctgggctgcg ggcttcatta gccatgtgtc gctggtcgaa 1140
48 atcctgatcc cattctgtgt gatggcgatt gccaatggcg cgatctaccc tattgttgtc 1200
49 gcccaggcgc tgcgtccctt cccacacgca actggtcgcg ccgcagcgtt gcagaacact 1260
50 cttcaactgg gtctgtgctt cctcgcaagt ctggtagttt cctggctgat cagtatcagc 1320
51 acgccattgc tcaccaccac cagcgtgatg ttatcaacag taatgctggt cgcgctgggt 1380
52 tacatgatgc aacgttgtga agaagttggc tgccagaatc atggcaatgc cgaagtcgct 1440
53 catagcgaat cacactgacc tatatcgata tacttatact taggctgcta acaaaatttt 1500
```

```
54 gttgtatgat tgaaattagc ggcctatact aatttcgagt tgttaaagct acgataaata 1560
55 ttatgttttt acggggacag gatcgttccc gactcactat ggatagtcat ttcggcaagg 1620
56 gttcctcctt tccctctgtt ctacgtcgga ttatagactc gcggtttttt ctgcgagatt 1680
57 tctcacaaag cccaaaaagc gtctacgctg ttttaaggtt ctgatcaccg accagtgatg 1740
58 gagaaactat gagttcatcg tgtatagaag aagtcagtgt accggatgac aactggtacc 1800
59 gtatcgccaa cgaattactt agccgtgccg gtatagccat taacggttct gccccggcgg 1860
60 atattcgtgt gaaaaacccc gattttttta aacgcgttct gcaagaaggc tctttggggt 1920
61 taggcgaaag ttatatggat ggctggtggg aatgtgaccg actggatatg ttttttagca 1980
62 aagtcttacg cgcaggtctc gagaaccaac tcccccatca tttcaaagac acgctgcgta 2040
63 ttgccggcgc tcgtctcttc aatctgcaga gtaaaaaacg tgcctggata gtcggcaaag 2100
64 agcattacga tttgggtaat gacttgttca gccgcatgct tgatcccttc atgcaatatt 2160
65 cctgcgctta ctggaaagat gccgataatc tggaatctgc ccagcaggcg aagctcaaaa 2220
66 tgatttgtga aaaattgcag ttaaaaccag ggatgcgcgt actggatatt ggctgcggct 2280
67 ggggcggact ggcacactac atggcatcta attatgacgt aagcgtggtg ggcgtcacca 2340
68 tttctgccga acagcaaaaa atggctcagg aacgctgtga aggcctggat gtcaccattt 2400
69 tgctgcaaga ttatcgtgac ctgaacgacc agtttgatcg tattgtttct gtggggatgt 2460
70 tcgagcacgt cggaccgaaa aattacgata cctattttgc ggtggtggat cgtaatttga 2520
71 aaccggaagg catattcctg ctccatacta tcggttcgaa aaaaaccgat ctgaatgttg 2580
72 atccctggat taataaatat atttttccga acggttgcct gccctctgta cgccagattg 2640
73 ctcagtccag cgaaccccac tttgtgatgg aagactggca taacttcggt gctgattacg 2700
74 atactacgtt gatggcgtgg tatgaacgat tcctcgccgc atggccagaa attgcggata 2760
75 actatagtga acgctttaaa cgaatgttta cctattatct gaatgcctgt gcaggtgctt 2820
76 tccgcgcccg tgatattcag ctctggcagg tcgtgttctc acgcggtgtt gaaaacggcc 2880
77 ttcgagtggc tcgctaaagg ctattctatc gcccctctc cggggcgat ttcagatcag 2940
78 gcttctgtgc ctggttgatt catggcattt tctcgtgccg ccagcacacg ttctaccgta 3000
79 tctaccactg cctgagtttg tggatcgatt tcaatgttga cgcgtgcgcc aagttttttc 3060
80 ttcccaagag tcgtgcgttc cagtgtttcc ggaattaaat ggacgcaaaa acgcgttggc 3120
81 gtgacttcgc cgacggtcag gctaataccg tcgatgccaa taaatccttt gtacagaata 3180
82 tatttcatca actgactatc ctggacttta aaccagatct ggcgattatt ttctgaggtt 3240
83 aatattttcg ccacttcagc agtggtcata atatgacctg acattaagtg tccgccaatt 3300
84 tcatcactga atttcgccgc acgctcaacg tttacccaat cccccacttt taaatcgcca 3360
85 agattggtaa tgcgtaacgt ttctttcatc aggtcaaaac tgacatggtt gccgttaatt 3420
86 tccgtcacgg tcaggcagca accgttatgc gccacggaag caccggtttc caggccgtcc 3480
87 agcatgtggt cgggtaactc caccacatgc gtacgaaaat ttggtttctc gtcaatcgac 3540
88 accagttttg cggtgccctg tacaatcccc gtaaacatac ttacaactcc tgaaatcagt 3600
89 taagacattc tgttcagcac aatagcaggt ggaaaacgcc cttaccagtg aaggggtaag 3660
90 aatggctatt ttttcactgg agaattaata aatcctcgct acaatagact gaatttcccc 3720
91 tgcttcttct ttttgctgcc cattcaggcg gctttttagt ctctcatata actacaaata 3780
92 aaaggtgttc acgtgcagaa gtatatcagt gaagcgcgtc tgttattagc attagcaatc 3840
93 ccggtgattc tcgcgcaaat cgcccaaact gcgatgggtt ttgtcagtac cgtgatggcg 3900
96 <210> SEQ ID NO: 2
97 <211> LENGTH: 1212
98 <212> TYPE: DNA
99 <213> ORGANISM: Escherichia coli
101 <400> SEQUENCE: 2
102 atgcaacctg ggaaaagatt tttagtctgg ctggcgggtt tgagcgtact cggttttctg 60
103 gcaaccgata tgtatctgcc tgctttcgcc gccatacagg ccgacctgca aacgcctgcg 120
104 tctgctgtca gtgccagcct tagtctgttc cttgccggtt ttgccgcagc ccagcttctg 180
105 tggggccgc tctccgaccg ttatggtcgt aaaccggtat tattaatcgg cctgacaatt 240
```

RAW SEQUENCE LISTING   DATE: 08/04/2004
PATENT APPLICATION: US/09/805,681B   TIME: 09:09:50

Input Set : A:\1137r00.app
Output Set: N:\CRF4\08042004\I805681B.raw

```
106 tttgcgttag gtagtctggg gatgctgtgg gtagaaaacg ccgctacgct gctggtattg 300
107 cgttttgtac aggctgtggg tgtctgcgcc gcggcggtta tctggcaagc attagtgaca 360
108 gattattatc cttcacagaa agttaaccgt atttttgcgg ccatcatgcc gctggtgggt 420
109 ctatctccgg cactggctcc tctgttagga agctggctgc tggtccattt ttcctggcag 480
110 gcgattttcg ccaccctgtt tgccattacc gtggtgctga ttctgcctat tttctggctc 540
111 aaacccacga cgaaggcccg taacaatagt caggatggtc tgacctttac cgacctgcta 600
112 cgttctaaaa cctatcgcgg caacgtgctg atatacgcag cctgttcagc cagttttttt 660
113 gcatggctga ccggttcacc gttcatcctt agtgaaatgg gctacagccc ggcagttatt 720
114 ggtttaagtt atgtcccgca aactatcgcg tttctgattg gtggttatgg ctgtcgcgcc 780
115 gcgctgcaga aatggcaagg caagcagtta ttaccgtggt tgctggtgct gtttgctgtc 840
116 agcgtcattg cgacctgggc tgcgggcttc attagccatg tgtcgctggt cgaaatcctg 900
117 atcccattct gtgtgatggc gattgccaat ggcgcgatct accctattgt tgtcgcccag 960
118 gcgctgcgtc ccttcccaca cgcaactggt cgcgccgcag cgttgcagaa cactcttcaa 1020
119 ctgggtctgt gcttcctcgc aagtctggta gtttcctggc tgatcagtat cagcacgcca 1080
120 ttgctcacca ccaccagcgt gatgttatca acagtaatgc tggtcgcgct gggttacatg 1140
121 atgcaacgtt gtgaagaagt tggctgccag aatcatggca atgccgaagt cgctcatagc 1200
122 gaatcacact ga                                                     1212
125 <210> SEQ ID NO: 3
126 <211> LENGTH: 389
127 <212> TYPE: PRT
128 <213> ORGANISM: Escherichia coli
130 <400> SEQUENCE: 3
131 Met Lys Ile Asn Tyr Pro Leu Leu Ala Leu Ala Ile Gly Ala Phe Gly
132   1               5                  10                  15
134 Ile Gly Thr Thr Glu Phe Ser Pro Met Gly Leu Leu Pro Val Ile Ala
135              20                  25                  30
137 Arg Gly Val Asp Val Ser Ile Pro Ala Ala Gly Met Leu Ile Ser Ala
138          35                  40                  45
140 Tyr Ala Val Gly Val Met Val Gly Ala Pro Leu Met Thr Leu Leu Leu
141      50                  55                  60
143 Ser His Arg Ala Arg Arg Ser Ala Leu Ile Phe Leu Met Ala Ile Phe
144  65                  70                  75                  80
146 Thr Leu Gly Asn Val Leu Ser Ala Ile Ala Pro Asp Tyr Met Thr Leu
147              85                  90                  95
149 Met Leu Ser Arg Ile Leu Thr Ser Leu Asn His Gly Ala Phe Phe Gly
150              100                 105                 110
152 Leu Gly Ser Val Val Ala Ala Ser Val Val Pro Lys His Lys Gln Ala
153          115                 120                 125
155 Ser Ala Val Ala Thr Met Phe Met Gly Leu Thr Leu Ala Asn Ile Gly
156      130                 135                 140
158 Gly Val Pro Ala Ala Thr Trp Leu Gly Glu Thr Ile Gly Trp Arg Met
159 145                 150                 155                 160
161 Ser Phe Leu Ala Thr Ala Gly Leu Gly Val Ile Ser Met Val Ser Leu
162              165                 170                 175
164 Phe Phe Ser Leu Pro Lys Gly Gly Ala Gly Ala Arg Pro Glu Val Lys
165              180                 185                 190
167 Lys Glu Leu Ala Val Leu Met Arg Pro Gln Val Leu Ser Ala Leu Leu
168          195                 200                 205
170 Thr Thr Val Leu Gly Ala Gly Ala Met Phe Thr Leu Tyr Thr Tyr Ile
```

RAW SEQUENCE LISTING   DATE: 08/04/2004
PATENT APPLICATION: US/09/805,681B   TIME: 09:09:50

Input Set : A:\1137r00.app
Output Set: N:\CRF4\08042004\I805681B.raw

```
171    210                215                220
173 Ser Pro Val Leu Gln Ser Ile Thr His Ala Thr Pro Val Phe Val Thr
174 225                230                235                240
176 Ala Met Leu Val Leu Ile Gly Val Gly Phe Ser Ile Gly Asn Tyr Leu
177                245                250                255
179 Gly Gly Lys Leu Ala Asp Arg Ser Val Asn Gly Thr Leu Lys Gly Phe
180                260                265                270
182 Leu Leu Leu Leu Met Val Ile Met Leu Ala Ile Pro Phe Leu Ala Arg
183        275                280                285
185 Asn Glu Phe Gly Ala Ala Ile Ser Met Val Val Trp Gly Ala Ala Thr
186        290                295                300
188 Phe Ala Val Val Pro Pro Leu Gln Met Arg Val Met Arg Val Ala Ser
189 305                310                315                320
191 Glu Ala Pro Gly Leu Ser Ser Ser Val Asn Ile Gly Ala Phe Asn Leu
192                325                330                335
194 Gly Asn Ala Leu Gly Ala Ala Ala Gly Gly Ala Val Ile Ser Ala Gly
195                340                345                350
197 Leu Gly Tyr Ser Phe Val Pro Val Met Gly Ala Ile Val Ala Gly Leu
198        355                360                365
200 Ala Leu Leu Leu Val Phe Met Ser Ala Arg Lys Gln Pro Glu Thr Val
201        370                375                380
203 Cys Val Ala Asn Ser
204 385
207 <210> SEQ ID NO: 4
208 <211> LENGTH: 391
209 <212> TYPE: PRT
210 <213> ORGANISM: Rhodococcus fascians
212 <400> SEQUENCE: 4
213 Met Pro Phe Ala Ile Tyr Val Leu Gly Ile Ala Val Phe Ala Gln Gly
214  1                 5                10                15
216 Thr Ser Glu Phe Met Leu Ser Gly Leu Ile Pro Asp Met Ala Gln Asp
217                20                25                30
219 Leu Gln Val Ser Val Pro Thr Ala Gly Leu Leu Thr Ser Ala Phe Ala
220        35                40                45
222 Ile Gly Met Ile Ile Gly Ala Pro Leu Met Ala Ile Val Ser Met Arg
223        50                55                60
225 Trp Gln Arg Arg Arg Ala Leu Leu Thr Phe Leu Ile Thr Phe Met Val
226 65                70                75                80
228 Val His Val Ile Gly Ala Leu Thr Asp Ser Phe Gly Val Leu Leu Val
229                85                90                95
231 Thr Arg Ile Val Gly Ala Leu Ala Asn Ala Gly Phe Leu Ala Val Ala
232                100                105                110
234 Leu Gly Ala Ala Met Ser Met Val Pro Ala Asp Met Lys Gly Arg Ala
235        115                120                125
237 Thr Ser Val Leu Leu Gly Gly Val Ile Ile Ala Cys Val Val Gly Val
238        130                135                140
240 Pro Gly Gly Ala Leu Leu Gly Glu Leu Trp Gly Trp Arg Ala Ser Phe
241 145                150                155                160
243 Trp Glu Val Val Leu Ile Ser Ala Pro Ala Val Ala Ala Ile Met Ala
```

```
RAW SEQUENCE LISTING                    DATE: 08/04/2004
PATENT APPLICATION:  US/09/805,681B     TIME: 09:09:50

Input Set  : A:\1137r00.app
Output Set: N:\CRF4\08042004\I805681B.raw
```

```
244                 165             170             175
246 Ser Thr Pro Ala Asp Ser Pro Thr Asp Ser Val Pro Asn Ala Thr Arg
247             180             185             190
249 Glu Leu Ser Ser Leu Arg Gln Arg Lys Leu Gln Leu Ile Leu Val Leu
250             195             200             205
252 Gly Ala Leu Ile Asn Gly Ala Thr Phe Cys Ser Phe Thr Tyr Leu Ala
253     210             215             220
255 Pro Thr Leu Thr Asp Val Ala Gly Phe Asp Ser Arg Trp Ile Pro Leu
256 225             230             235             240
258 Leu Leu Gly Leu Phe Gly Leu Gly Ser Phe Ile Gly Val Ser Val Gly
259             245             250             255
261 Gly Arg Leu Ala Asp Thr Arg Pro Phe Gln Leu Leu Val Ala Gly Ser
262             260             265             270
264 Ala Ala Leu Leu Val Gly Trp Ile Val Phe Ala Ile Thr Ala Ser His
265         275             280             285
267 Pro Val Val Thr Leu Val Met Leu Phe Val Gln Gly Thr Leu Ser Phe
268     290             295             300
270 Ala Val Gly Ser Thr Leu Ile Ser Arg Val Leu Tyr Val Ala Asp Gly
271 305             310             315             320
273 Ala Pro Thr Leu Gly Gly Ser Phe Ala Thr Ala Ala Phe Asn Val Gly
274             325             330             335
276 Ala Ala Leu Gly Pro Ala Leu Gly Gly Val Ala Ile Gly Ile Gly Met
277             340             345             350
279 Gly Tyr Arg Ala Pro Leu Trp Thr Ser Ala Ala Leu Val Ala Leu Ala
280             355             360             365
282 Ile Val Ile Gly Ala Ala Thr Trp Thr Arg Trp Arg Glu Pro Arg Pro
283     370             375             380
285 Ala Leu Asp Thr Val Pro Pro
286 385             390
289 <210> SEQ ID NO: 5
290 <211> LENGTH: 391
291 <212> TYPE: PRT
292 <213> ORGANISM: Rhodococcus erythropolis
294 <400> SEQUENCE: 5
295 Met Pro Phe Ala Ile Tyr Val Leu Gly Leu Ala Val Phe Ala Gln Gly
296   1             5               10              15
298 Thr Ser Glu Phe Met Leu Ser Gly Leu Ile Pro Asp Met Ala Arg Asp
299             20              25              30
301 Leu Gly Val Ser Val Pro Ala Ala Gly Leu Leu Thr Ser Ala Phe Ala
302         35              40              45
304 Val Gly Met Ile Ile Gly Ala Pro Leu Met Ala Ile Ala Ser Met Arg
305     50              55              60
307 Trp Pro Arg Arg Arg Ala Leu Leu Thr Phe Leu Ile Thr Phe Met Leu
308 65              70              75              80
310 Val His Val Ile Gly Ala Leu Thr Ser Ser Phe Glu Val Leu Leu Val
311             85              90              95
313 Thr Arg Ile Val Gly Ala Leu Ala Asn Ala Gly Phe Leu Ala Val Ala
314             100             105             110
316 Leu Gly Ala Ala Met Ala Met Val Pro Ala Asp Met Lys Gly Arg Ala
```

VERIFICATION SUMMARY   DATE: 08/04/2004
PATENT APPLICATION: **US/09/805,681B**   TIME: 09:09:51

Input Set : **A:\1137r00.app**
Output Set: **N:\CRF4\08042004\I805681B.raw**